2018-01-01

From Babies to Robots: The Contribution of Developmental Robotics to Developmental Psychology

Cangelosi, A

http://hdl.handle.net/10026.1/11191

10.1111/cdep.12282
Child Development Perspectives

From Babies to Robots: The Contribution of Developmental Robotics to Developmental Psychology

Angelo Cangelosi, Matthew Schlesinger

Abstract
The latest developments in AI and machine learning, and the parallel advances in robotics, have very recently contributed to a shift in the approach to modeling human intelligence. These innovations, accompanied by the new emphasis on embodied and grounded cognition in AI and psychology, have led to the establishment of the field of Developmental Robotics. This is the interdisciplinary approach, built on the close collaboration of the disciplines of cognitive robotics and child psychology, to the autonomous design of behavioral and cognitive capabilities in artificial cognitive agents, such as robots, which takes direct inspiration from the developmental principles and mechanisms observed in children. We illustrate the benefits of this approach by presenting a detailed baby robot case study of the role of embodiment during early word learning, as well as an overview of several developmental robotics model of perceptual, social and language development.

Introduction
Computational models of cognition have significantly contributed to the definition, testing, and validation of psychology and neuroscience theories, including developmental psychology. Such computational approaches, ranging from symbolic rule-based systems, connectionist neural networks, and Bayesian models, have typically resulted from scientific and technological developments in artificial intelligence (AI) and its attempt to reproduce and simulate the uniqueness and complexity of human-like *adult symbolic* intelligence. However, since the origins of AI, there have been proposals to study the full spectrum of child development, rather than adult-like intelligence. This is for example what Alan Turing proposed in 1950:

> *"Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain." Turing (1, page 440).*

The latest developments in AI and machine learning, and the parallel advances in robotics, have contributed to a shift in the approach to modeling human intelligence. These innovations have been accompanied by an increased emphasis on embodied and grounded cognition in AI (2) and psychology (3,4). This has permitted the first attempts to realize Turing's vision, i.e. the idea that an embodied agent (e.g. robot), using a set of intrinsic motivation principles regulating the real-time interaction between its body, brain and environment, can autonomously acquire and develop an increasingly complex set of sensorimotor and mental capabilities, grounded in the interaction with its environment.

The field of Developmental Robotics (DR, hereafter) specifically aims to design sensorimotor and cognitive capabilities in (baby) robots by taking inspiration from child psychology and via the modelling of incremental, developmental changes. Thus DR relies on a highly interdisciplinary effort of empirical developmental sciences, including developmental psychology, neuroscience and comparative psychology, and on computational and robotics disciplines, such as robotics, machine learning and artificial intelligence (5,6). The history of developmental robotics can be traced back to the early 2000s, where the close collaboration of cognitive modelers and developmental psychologists led to the first, pioneering robotics models of development (7-9).

DR builds on, but further extends, the existing computational approaches of development. In classical computational models of developmental mechanisms, as in the seminal Klahr and Wallace (10) information processing model of Piagetian concrete operation tasks, the modeler has to define in detail very abstract, rule-like representations involved in the phenomenon studied. Even in classical connectionist models of language learning, such as the well known past tense simulations (11), or more recent large scale neural models of language learning (e.g. 12) the modeler has to provide a formal and pre-defined representation of the input stimuli features, though the neural network is capable to simulate qualitative changes in the learning pathway. DR models, instead, permit the exploitation of the role of sensorimotor knowledge, of motivational and attentional mechanisms and of interaction and environmental factors, in the acquisition of cognitive skills (see for further discussion: 5, 13, 14).

Direct modeling of developmental psychology experiments and data

One of the key aims of developmental robotics is to take explicit inspiration from human developmental mechanisms to design cognitive skills in robots. We can distinguish two main approaches to the handling of the relation between developmental psychology and developmental robotics. In the first case on the direct modelling of developmental psychology studies, the robot experiments are directly constructed to replicate specific child psychology experiments. This often permits the direct comparison (qualitative and/or quantitative) and prediction of empirical and modeling results. The other approach concerns a more generic, higher-level cognitively-inspired link between the broad developmental mechanism studied in child experiments and the general developmental aspects of the robotic algorithms. This is not a rigid separation of DR studies, as robotics models mostly lie in the continuum between the direct modelling and the bio-inspired approach.

This paper will concentrate on some prototypical examples of the direct modelling approach, since this better shows the benefits of a direct collaboration between developmental psychologists and roboticists. As we will show in the case study below on the Epigenetic Robotics Architecture (ERA) (15-17) for robot experiments on early word learning, the direct modelling of developmental psychology data can also lead to novel predictions on developmental phenomena, further validated by new child experiments. For a discussion of examples on the general, bio-inspired approach, see the extended analysis of the DR literature in Cangelosi and Schlesinger (5).

Other examples of the direct comparison of child and robot data have been proposed in the field of perceptual development. Schlesinger et al. (18) presented a model of perceptual completion directly simulating Amso and Johnson's (19) experiments on unity perception in young infants. This DR computational model directly compares the performance of the simulation model (simulated distribution of scans) with 3-month-old infants' eye tracking data. The model provides an operational explanation of the neural and developmental mechanisms in early perception.

In social interaction studies, Nagai and collaborators (20, 21) model the developmental, stage-like emergence of shared gaze. They explicitly follow Butterworth's (22) developmental framework on the incremental acquisition of different attentional gaze strategies: from the ecological stage (the infant looks at an interesting object regardless of the caregiver's gaze direction), to the geometric (joint attention only when the object is in the infant's field of view) and representational (the infant can find salient object outside its own field of view) gaze strategies stages.

A Case Study: Embodied Word Learning

A prototypical example of a DR study that aims to directly model child psychology data and to use the computational model to make predictions of language learning mechanisms is provided by Morse, Cangelosi, Smith, and collaborators (15, 16, 23). This DR model addresses the issue of embodiment factors, that is, how spatial locations, and their corresponding postural changes, play a key role in infants' word learning.

The robot's cognitive architecture

To model the role of embodiment in word learning, a DR modeling framework, called the Epigenetic Robotic Architecture (ERA) (15), has been proposed. This cognitive architecture is based on an ensemble of artificial neural networks used to implement the learning from multimodal stimuli (visual, speech, postural) and to control the robot's behavior (Figure 1). The DR architecture consists of multiple maps, each realized via a Self-Organizing Map (SOM) (24). A SOM is an artificial neural network where the output layer consists of a 2-dimensional grid of neuron (called map). After training, the output neurons self-organize to create a similarity map. The maps used in the ERA architecture (color map, shape map, postural map) have been pre-trained respectively to build a categorical similarity representation of color stimuli, object shapes and the robot's own body posture. Another map (speech map) has been hardwired to encodes word representations. The map-to-map links are constituted by Hebbian connections. These connections implement the associative learning between the most active node in each map, activated by visual, postural and speech stimuli.

< figure 1 about here >

The organization and properties of such a robot architecture are purposely chosen to operationalize the key developmental principles and mechanisms needed for language development. The SOMs constitute the building block of a hierarchical set of interconnected cortical brain areas. Specifically, the use of pre-trained SOMs has the purpose to endow the robot with the pre-linguistic capability to recognize and categorize object’s colors and shapes, as well as creating a homunculus-type body representation via a motor-babbling training stage (i.e. a pre-training stage where the robot randomly moves its limbs to allow it to learn a body representation map). The Hebbian connections between the active nodes in each map implement the principle of multimodal associate learning needed to link the name of an object to its visual features (color and/or shape category) and to specific postures.

In addition to the neural network architecture, the robot is pre-programmed with an intrinsic motivation mechanism to gaze at the position of moving entities, such as when waving a hand or shaking an object. This implements the developmental principle that infants have a tendency to pay attention to moving objects. It can be exploited by the experimenter to make sure that the robot’s visual system focuses on the object in sight, or towards the spatial location where the hand is moved, when the name of the target object is uttered during the language learning experiment.

The baby humanoid robot iCub is used for the DR experiments (Figure 1b). The iCub is an open source robotic platform recently developed as a benchmark platform for cognitive and developmental robotics experiments (25). The use of the robot, controlled by its neural cognitive architecture with associative connections trained online during the experiment, further implements the developmental principle that word-object associations are the direct result of the robot’s interaction with its tutor and its physical environment.

< figure 1 about here >

The Baldwin task

The robot experimental procedure follows exactly the one used in child psychology experiments on early word learning, i.e. the Baldwin Task used in the Samuelson et al. study (26). The experimenter sits in front of the robot, with a white table where objects will be shown and labeled. Every time an object is shown, the robot shifts its posture (torso, arms and gaze) to look at the object, and learns to categorize it according to its visual features (e.g. color and shape). Each trial consists of nine steps:

Steps 1-2	The experimenter starts by showing two novel objects to the robot: (i) the target object, whose name has to be learned, and (ii) the foil object acting as distractor. These objects are shown one at a time respectively on the left and right location of the table.
Steps 3-4	The two objects are shown again
Steps 5	The experimenter hides the objects, directs the robot's attention towards the right side where the first (target) object was shown and says: "This is a Modi".
Steps 6-7	The two objects are shown again, one at a time, as in the initial steps.
Step 9	Both objects are presented simultaneously, in a new location at the center of the table, and the robot is asked "Find the Modi".

The robot experiments

In Morse et al. (16), a combination of 5 experiments with robot participants and 4 with children are carried out. These use the default Baldwin object-label mapping tasks in which names are either encountered in the absence of their target, or an Interference task (i.e. when their target is present, but in a location previously associated with a foil). The first set of iCub robot experiments was used to replicate and extend the original paradigm by Samuelson et al. (26). The robot experiments first replicated existing infant tasks and data: (Exp.1) the default Baldwin Task, (Exp.2) a Switch Task in which the position of the two objects in steps 3-4 is swapped, to stop the object-name association, and (Exp.3) the Interference Task. Two extra, novel robot experiments test the effects of a second postural change (sitting/standing) in addition to the left/right posture change of the previous experiments: (Exp.4) Posture Change Task, when in step 5 the robot changes position from sitting to standing, and (Exp.5) the Interference Posture Change Task, which follows the Interference task, but with the same sitting-to-standing posture change in the naming test at the final step. The results of these novel robot experiments demonstrated that in the Posture Change Task the position change disrupts the object-name association and the iCub randomly picks any of the two objects. On the contrary, in the Interference Posture Change Task, when the posture shift was instituted also during the naming event of the final step, the robot learns and maintains the association between the target object and its "Modi" name.

The child experiments

Four child experiments were also carried out, with the aim of replicating with infants four of the robot experiments: Baldwin Task, Interference Task, Posture Change Task and Interference with Posture Change Task. The results of the experiments with children replicated the same pattern of results from the robot studies. In particular, the child data validated the novel robot modeling results of both the tasks with posture change, which had not been previously studied in child psychology. This showed that despite spatial location being task-irrelevant in the Interference with Posture Change Task, infants (as predicted with robots) use body-centric spatial contingency over temporal contingency to map the name to the object. Both infants and robots remember the name-object mapping even in new spatial locations. In addition, the analyses of the robot's neural control architecture show how this memory can emerge. The iCub study demonstrates an exquisite coupling of the body's momentary spatial orientation and internal cognitive operations (see also, 26). The iCub model suggests that word and object features are fundamentally tied to bodily information; by contrast, the model proposed by Samuelson et al. links word-object associations to an integrated spatial representation in a way that can, ultimately, generalize over space, while remaining fully coupled to the real-time orientation of the body in space.

Lesson learned

Overall, this model demonstrates that it is possible to build an embodied cognitive system that develops linguistic and sensorimotor capabilities through interactions with the world, closely resembling multiple child development phenomena (16). This is achieved through the design of a cognitive architecture implementing key developmental mechanisms and principles during early word learning stages: (i) the pre-linguistic capability to recognize and categories object features; (ii) capability to segment heard words; (iii) a representation of own's body posture with respect to its environment; (iv) an attentional mechanisms guided by motion perception; (v) the learning mechanism for the acquisition of multimodal associations during interaction with a tutor.

The interaction of such perceptual, linguistic and sensorimotor capabilities permits not only the replication of known child data, but also the prediction of novel phenomena involved in the postural biases in word learning. This approach can be extended to shed light on how children change their word learning abilities over the longer timescales of development. For example, the ERA cognitive architecture and learning principles have been extended to model other developmental phenomena, such as mutual exclusivity (17) and U-shape phenomena in Universal Phonetic Discrimination stage (27). The ERA developmental architecture has been shown to be particularly suitable to model qualitative changes in development. The interaction between learning mechanisms, the resulting embodied behavior of the agent, and the opportunities for learning that the environment provides, can account for the staged development of cognitive abilities. In the ERA architecture, two simple mechanisms account for the developmental transitions and the multiple early language learning phenomena replicated. One is the realization of 'neural readiness', i.e. the focus on changes in the neural substrate resulting from ongoing learning which facilitate the acquisition of new knowledge and skills. The second developmental phenomenon is 'perceptual readiness', i.e, the focus on the multimodal perceptual requirements supporting the learning of new tasks (see 23 for extended discussion on these two developmental mechanisms).

The study of multiple developmental phenomena is possible via the realization of DR experiments utilizing a physical robot agent, a cognitive architecture for multimodal associative learning linking own sensorimotor representation with external linguistic stimuli, and the result of direct interaction between the (robot) child and the (human) experimenter. All this permits the direct testing of hypotheses on postural and spatial biases in cognitive development, minimizing the role of the computational modeler in the detailed implementation of the features and structure of perceptual and linguistic representations, and in the results of the learning interactions.

Looking Ahead: Scientific Challenges and Applications

The detailed analysis of the iCub robot's experiments of the embodiment strategies of early word learning, and the multiple examples of DR models of perceptual, social and linguistic development in robots, directly grounded on child psychology studies, shows the multiple benefits of the baby robot approach. These achievements have set the bases for new scientific challenges in the field. Key themes for future work in DR include the focus on open-ended, cumulative learning, the modeling of physical and neural maturational mechanisms interacting with both evolutionary and ontogenetic development, and ethical aspects in child-robot interaction research (see 5 for full discussion of these challenges).

For example, the challenge on open-ended learning refers to the fact that the robot keeps learning in new interactions with its environment. Cumulative learning refers to the simultaneous and cumulative acquisition of cognitive skills. This challenge involves the idea of 'raising' an infant robot into early childhood, if not longer, in an artificial "robot kindergarten". In Araki et al. (28), a learning robot interacts for a full week with an experimenter, who teaches it the names of 200 objects during numerous online learning sessions. Adams et al. (29) have also proposed the approach of a "virtual school student" in the robot kindergarten/school. This method foresees the implementation of a virtual student robot growing in both "Preschool Learning" (based on open, long term interaction experiments of sensorimotor skills and basic cognitive capabilities) and "School Learning" environment (for long-term practice of higher cognitive abilities).

In addition to such scientific challenges, the DR models of cognitive development have important implications for current and future applications in child psychology, child rehabilitation and education. For example, several pioneering investigations have looked at the translation of robot modeling research, especially the studies on social interaction, into applications of social assistive robotics as for children with autism spectrum disorder (ASD) (e.g. 30, 31). Scassellati et al. (31) analyze the achievements in this field and suggest that these improved social skills and behaviors via robot interaction are the consequences of the fact that robots provide novel sensory stimuli to the ASD child. A similar approach to the use of social developmental robots as therapy for ASD children has been applied to other disabilities as in the treatment for children with diabetes (32) and with mobility and motor disabilities (33). Although standard, pre-programmed robotic algorithms can also be used in robot therapy for children with disabilities, DR models offer the advantage of being intrinsically focused on typical (and atypical) developmental changes.

Finally, a key area of development both for scientific and application-oriented work of DR is the use of robots for education. Karim and colleagues (34) have recently reviewed the potential for using robots for STEM education. They discuss how the many studies on robots' involvement for teaching and tutoring for disciplines such as mathematics, science, and language, have their roots in classical psychology socio-constructivist theories (e.g. 35) and in modern pedagogic theories on active learning (36).

Acknowledgements

The work of Angelo Cangelosi was supported by the EU H2020 Marie Skłodowska-Curie projects APRIL (674868) and DCOMM (676063).

References

1. Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59, no. 236, 433-60.
2. Brooks, R. A. (1990). Elephants Don't Play Chess. *Robotics and Autonomous Systems*, *6(1)*, 3-15.
3. Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology, 59*, 617-45.
4. Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M.H., McRae, K., Spivey, M. (2013). Computational grounded cognition: A new alliance between grounded cognition and computational modelling. *Frontiers in Psychology*, *6*, 612, 1-11. doi:10.3389/fpsyg.2012.00612
5. Cangelosi, A., & Schlesinger, M. (2015). *Developmental Robotics: From Babies to Robots*. Cambridge, MA: MIT Press
6. Lungarella, M., Metta, G., Pfeifer, R. & Sandini G. (2003). Developmental robotics: A survey, *Connection Science*, *15*(4), 151–190.
7. Weng, J. Y., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M., & Thelen, E. (2001). Artificial Intelligence - Autonomous Mental Development by Robots and Animals. *Science, 291*, no. 5504, 599-600.
8. Balkenius, C., Zlatev, J., Kozima, H., Dautenhahn, K., & Breazeal, C., (2001). *First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, Lund, Sweden, 2001.
9. Asada, M., MacDorman, K. F., Ishiguro, H., & Kuniyoshi Y. (2001). Cognitive Developmental Robotics: As a New Paradigm for the Design of Humanoid Robots. *Robotics and Autonomous Systems*, *37*, 185-193.
10. Klahr, D., & Wallace, J. G. (1970). An information processing analysis of some Piagetian experimental tasks. *Cognitive Psychology*, *1*, 358-387.

11. Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In D.E. Rumelhart et al. (Eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press
12. Golosio, B., Cangelosi, A., Gamotina, O., & Masala, G.L. (2015). A cognitive neural architecture able to learn and communicate through natural language. *PLoS ONE*, *10(11)* 10.1371/journal.pone.0140866
13. Oudeyer, P.-Y. (2017), What do we learn about development from baby robots?. *WIREs Cognitive Science*, 8: e1395. doi:10.1002/wcs.1395
14. Merrick, K. (2017). Value Systems for Developmental Cognitive Robotics: A Survey, *Cognitive Systems Research*, *41*, p38 – 55.
15. Morse, A. F., DeGreeff, J., Belpeame, T., & Cangelosi, A. (2010). Epigenetic Robotics Architecture (ERA*). IEEE Transactions on Autonomous Mental Development*, *2(4)*, 325-339.
16. Morse, A., Belpaeme, T., Smith, L., Cangelosi, A. (2015). Posture affects how robots and infants map words to objects. *PLoS ONE*, *10*(3) 10.1371/journal.pone.0116012
17. Twomey, K. E., Morse, A. F., Cangelosi, A., & Horst, J. (2016). Children's referent selection and word learning: Insights from a developmental robotic system. *Interaction Studies*, 17(1), 101–127 10.1075/is.17.1.05two
18. Schlesinger, M., Amso, D., & Johnson, S. P. (2007). Simulating Infants' Gaze Patterns During the Development of Perceptual Completion. *Seventh International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems* Lund, Sweden.
19. Amso, D., & Johnson, S. P. (2006). Learning by Selection: Visual Search and Object Perception in Young Infants. *Developmental Psychology*, *42*(6), 1236-45.
20. Nagai, Y., Hosoda, K., Morita, A., & Asada, M. (2003). A Constructive Model for the Development of Joint Attention. *Connection Science*, 15(4), 211-29.
21. Nagai, Y., Asada, M., Hosoda, K. (2006). Learning for joint attention helped by functional development. *Advanced Robotics*, *20*(10), 1165-1181.
22. Butterworth, G. (1991). The Ontogeny and Phylogeny of Joint Visual Attention. In A. Whiten *(Ed.), Natural Theories of Mind*, 223-32. Cambridge, MA: Blackwell, 1991.
23. Morse, A., & Cangelosi, A. (2017). Why are there developmental stages in language learning? A developmental robotics model of language development. *Cognitive Science*, 10.1111/cogs.12390
24. Kohonen, T. (1982). Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics, 43*(1), 59–69.
25. Metta, G., Natale, L., Nori, F., Sandini, G., Vernon, D., Fadiga, L. von Hofsten C., Rosander K., Lopes M., Santos-Victor J., Bernardino A., & Montesano L. (2010). The iCub humanoid robot: An open-systems platform for research in cognitive development. *Neural Networks*, *23*(8), 1125-1134.
26. Samuelson, L. K., Smith, L. B., Perry, L. K., Spencer, J. P. (2011). Grounding word learning in space. *PLoS ONE*, 6, e28095.
27. Morse, A. F., Belpaeme, T., Cangelosi, A., & Floccia, C. (2011). Modeling U-Shaped Performance Curves in Ongoing Development. *23rd Annual Meeting of the Cognitive Science Society*, Austin.

28. Araki, T., Nakamura, T., Nagai T. (2013). Long-Term Learning of Concept and Word by Robots: Interactive Learning Framework and Preliminary Results. *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Tokyo Big Sight, Tokyo, Japan, November 3-7
29. Adams, S. S., Arel, I., Bach, J., R. Coop, Furlan, R., Goertzel, Hall, J. S., Samsonovich, A., Scheutz, M., Schlesinger, A., Shapiro, S. C., & Sowa, J. F. (2012). Mapping the Landscape of Human-Level Artificial General Intelligence. *AI Magazine*, 33(1), 25-41.
30. Dautenhahn, K. (1999). Robots as Social Actors: Aurora and the Case of Autism." *Proc. CT99, The Third International Cognitive Technology Conference*, August, San Francisco.
31. Scassellati, B., Admoni, H., & Matarić, M. (2012). Robots for Use in Autism Research. *Annual Review of Biomedical Engineering*, 14, 275-94.
32. Belpaeme, T., Baxter, P., Read, R., Wood, R., Cuayáhuitl, H., Kiefer, B., Racioppa, S., Kruijff-Korbayová, I., Athanasopoulos, G., Enescu, V., Looije, R., Neerincx, M., Demiris, Y., Ros-Espinoza, R., Beck, A., Cañamero, L., Hiolle, A., Lewis, M., Baroni, I., Nalin, M., Cosi, P., Paci, G., Tesser, F., Sommavilla, G., & Humbert, R. (2012) Multimodal Child-Robot Interaction: Building Social Bonds. *Journal of Human-Robot Interaction*, 1(2), 33-53.
33. Sarabia, M., & Demiris, Y. (2013). A Humanoid Robot Companion for Wheelchair Users. In *Social Robotics*, edited by G. Herrmann et al. Lecture Notes in Computer Science, 432-41. Berlin, Germany: Springer
34. Karim, M. E., Lemaignan, S., & Mondada, F. (2015). A review: Can robots reshape K-12 STEM education?. *In Advanced Robotics and its Social Impacts (ARSO), 2015 IEEE International Workshop on* (pp. 1-8). IEEE.
35. Vygotsky, L. S. (1980). Mind in society: The development of higher psychological processes. Harvard University Press
36. Harmin, M., & Toth, M. (2006). *Inspiring active learning: A complete handbook for today's teachers*. ASCD, 2006.

This is the author's accepted manuscript. The final published version of this work is published by John Wiley in Child Development Perspectives available at DOI 10.1111/cdep.12282. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

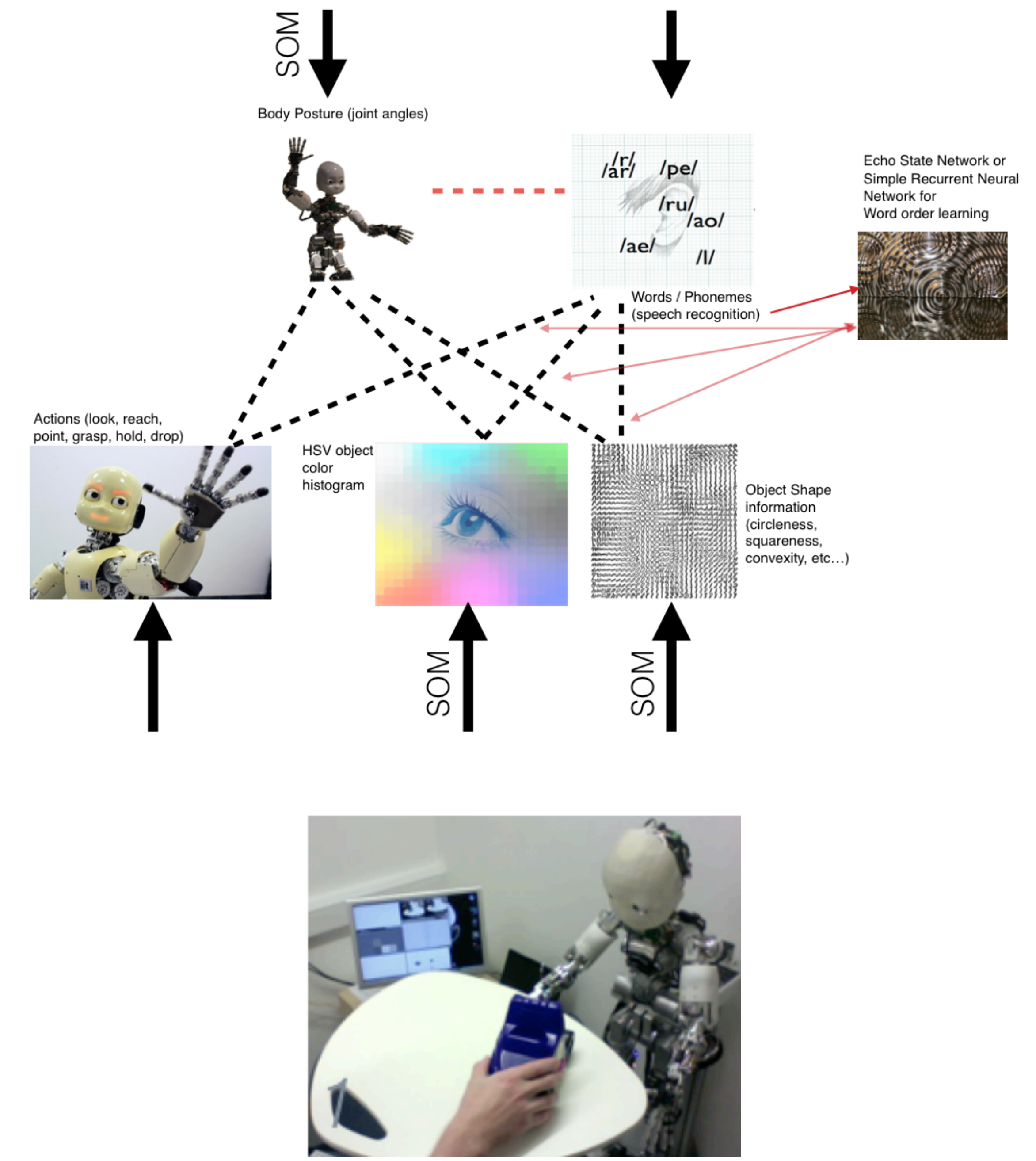

Figure 1 Cognitive architecture (top) and experimental setup (bottom) of the iCub model of embodied word learning (from Morse et al. 2015).